Computation of Singularities in B.V.P.'S Using Finite Differences with Block-Elimination

by

Yaqoub Mustafa Shehadeh

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

MATHEMATICS

June, 1992

Order Number 1354099

Computation of singularities in B.V.P.'s using finite differences with block-elimination

Shehadeh, Yaqoub Mustafa, M.S.

King Fahd University of Petroleum and Minerals (Saudi Arabia), 1992

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

COMPUTATION OF SINGULARITIES IN B.V.P.'S USING FINITE DIFFERENCES WITH BLOCK-ELIMINATION

BY

YAQOUB MUSTAFA SHEHADEH

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

In

MATHEMATICS

JUNE, 1992

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

COLLEGE OF GRADUATE STUDIES

This thesis, written by *YAQOUB MUSTAFA SHEHADEH* under the direction of his Thesis Advisor and approved by his Thesis Committee, has been presented to and accepted by the Dean of the College of Graduate Studies, in partial fulfillment of the requirements for the degree of *MASTER OF SCIENCE* in *MATHEMATICS*.

THESIS COMMITTEE

June 30, 1992

Dr. Basem Attili (Chairman)

June 30, 1992

Dr. Mahmoud Sarhan (Member)

June 30, 1992

Dr. Isik Tarman (Member)

30-6-92

Department Chairman

Dean, College of Graduate Studies

Date: 1/7/92

To My beloved parents, sisters and brothers

ACKNOWLEDGEMENT

All praise be to "ALLAH", the lord of the world, the Almighty, with whose gracious help it was possible to accomplish this work. May peace and blessing be upon Mohammad the last of the Messengers.

I would like to express my deep appreciation and gratitude to my supervisor, Dr. Basem S. Attili, for his patient guidance and his generous support and encouragement. I also would like to thank the other members of my thesis committee, Dr. Mahmoud Sarhan and Dr. Isik Tarman. I would like to acknowledge with great thanks to Dr. Mohammed Al-Bar, the Chairman of the Department of Math.

Special thanks are due to Dr. M. Z. Abu-Sbeih, Mr. Muhammad Awad and Mr. Abdallah Shuibi, for their assistance in producing this work in a nice form and to all my brothers, friends and many others for the memorable days we shared together.

I am also thankful to Mr. Saadat Ali Siddiqui for doing an excellent job in typing this thesis.

الخلاصة

العنوان :- الحساب العددي لنقاط حرجة لمعادلات تفاضلية من الدرجة الثانية باستخدام طرقة الفروق للحدونة العددية مع الحذف الجزئي.

الاسم :- يعقوب مصطفى شحاده

التخصص :- الرياضيات

التاريخ :- حزيران ١٩٩٢م

في هذه الرسالة سنعرض طريقة مباشرة لحساب النقاط الحرجة البسيطة والتكعيبية ونقاط التشعب البسيطة لمعادلات تفاضلية من الدرجة الثانية. سوف نستخدم طريقة الفروق المحدودة العددية لحل المعادلات التفاضلية . وبما ان الجزء الرئيسي من المصفوفات التي تنتج هو نفسه في الجميع فهذا سيمكننا من استخدام فكرة الحذف الجزئي. مما يقلل من العمليات الحسابية المطلوبة للقيام بالمهمة. وللحصول على نتائج ادق فسنستعمل تقريب ريتشارسون باستعمال خطوات ذات حجم صغير. سنعرض ايضاً امثله عددية لما تقدم ذكره.

درجة الماجستير في العلوم

جامعة الملك فهد للبترول والمعادن

الظهران - المملكةالعربية السعودية

حزيران ١٩٩٢م

ABSTRACT

Title : **COMPUTATION OF SINGULARITIES IN B.V.P.'S USING FINITE DIFFERENCES WITH BLOCK-ELIMINATION**

Name : Yaqoub Mustafa Shehadeh

Major Field : Mathematics

Date : June, 1992

In this thesis, we will consider the numerical computation of simple turning points, simple bifurcation points and cubic turning points, for two point boundary value problems. A direct method will be employed to characterize the B.V.P.'s. Since the main block of all systems involved is the same, this will allow us to use block elimination which will reduce the work by a large factor. To obtain accurate results we will use Richardson extrapolation with small sizes. Numerical examples will also be presented.

MASTER OF SCIENCE DEGREE

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS

DHAHRAN, SAUDI ARABIA

JUNE, 1992.

Contents

1 DEFINITIONS AND CLASSIFICATIONS 4

1.1 Definitions 4

1.2 Turning Points 6

1.3 Bifurcation Points 10

2 CHARACTERIZATION OF SIMPLE SINGULARITIES 14

2.1 Characterization of Simple (Quadratic) Turning Points . . . 15

2.2 Characterization of Simple Bifurcation Points 20

2.3 Characterization of Cubic Turning Points 25

3 NUMERICAL DETAILS 27

3.1 Discretization by Finite - Differences 28

3.2 Block Elimination 41

3.2.1 Algorithm 42

3.2.2 Algorithm 43

3.2.3 Algorithm 45

3.2.4 Algorithm 48

3.3 The Partials of g and $\tilde{g}$ 51

3.4 Choice of R and T 56

4 NUMERICAL EXAMPLES 58

INTRODUCTION

Many physical problems can be formulated as a parameter dependent nonlinear system of equations of the form

$$F(z, \lambda) = 0, \tag{1}$$

where F is a C^3 - functions which maps $R^n \times R$ into R^n. Such problems include systems of differential equations, integral equations or algebraic systems. We will assume that (1) is a finite difference approximation of a two point boundary value problem of the form

$$\begin{cases} z''(x) + f(x, z(x), \lambda) = 0 & a \leq x \leq b \\ z(a) = \alpha, \quad z(b) = \beta. \end{cases}$$

If (z_0, λ_0) is a regular point of (1); that is, $F_z(z_0, \lambda_0)$ is nonsingular, then the implicit function theorem insures the existence and uniqueness of a solution to (1). While if (z_0, λ_0) is a singular point of (1); that is,

$F_z(z_0, \lambda_0)$ is sigular, then all procedures which try to find a solution for (1) might fail or encounter difficulty close to such singular points.

We will consider three types of singular points, namely, simple turning points, simple bifurcation points and cubic turning points. Analytical aspects of turning points and bifurcation points of (1) have been discussed by many authors, see for example the review by Stakgold[16] . We will employ a direct method proposed by Griewank and Reddien[7] and developed by Attili[1,2,4] to characterize such singular points. The systems which result from such characterization will have the common main block F_z in all of them. This suggests the use of block- elimination to save on the number of operations needed to compute the solutions at each step. Simply, at each iteration we will compute the LU-factorization of F_z and use it to solve all the systems involved by just Back substitutions. The saving is due to the fact that the LU-factorization requires $O(N^3)$ operations while the back-substitution requires $O(N^2)$ operations.

For accuracy, we will solve smaller systems and make use of Richardson extrapolation.

In chapter 1, we will present the definitions of the three types of singularities. The characterization of such singularities will be discussed in chapter 2. The idea will be to extend the systems involved to obtain a larger system for which the singular point is an isolated solution. In chapter 3, we will present the details of the numerical procedure we will follow together with the block- elimination and Richardson and extrapolation. Finally, numerical examples and comparisons between solving the systems with and without block-elimination will be given in chapter 4 in terms of CPU time required.

Chapter 1

DEFINITIONS AND CLASSIFICATIONS

In this chapter we recall the basic concepts regarding turning points and bifurcation points.

1.1 Definitions

Consider the problem

$$f(x, \lambda) = 0, \tag{1. 1.1}$$

where f is a C^3-mapping from $R^n \times R \to R^n$. Let $(x, \lambda) = (x(s), \lambda(s))$ represent a solution arc, where s is a real parameter.

Definition 1.1.1 *$(x_0, \lambda_0) = (x(s_0), \lambda(s_0))$ is called a regular point if $f_x(x_0, \lambda_0)$ is nonsingular.*

The implicit function theorem ensures the existence and uniqueness of a solution branch in the nieghborhood of (x_0, λ_0) where (x_0, λ_0) is a regular point

A simple version of the implicit function theorem is stated below:

Theorem 1.1.1 *Let $f(x_0, \lambda_0) = 0$ where f maps $R \times R \to R$ and let f be continuously differentiable in some open region containing the point (x_0, λ_0) of the (x, λ) plane. Then, if $f_x(x_0, \lambda_0) \neq 0$, there exist $\alpha > 0$ and $\beta > 0$ such that the equation $f(x, \lambda) = 0$ has a unique solution $x = x(\lambda)$ when $x_0 - \alpha < x < x_0 + \alpha$ such that $\lambda_0 - \beta < \lambda < \lambda_0 + \beta$.*

For other versions of this theorem and for more details see Ioose [7].

If $f_x^0 = f_x(x_0, \lambda_0)$ is singular, (x_0, λ_0) is called a singular point and various types of behaviour can occur.

Assumptions

We assume:

(H_1) f_x^0 has a one-dimensional null space spanned by $\phi_0 \in R^n$; that is, $N(f_x^0) =$ span $\{\phi_0\}$.

(H_2) $f_x^{0^T}$ has a one-dimensional null space spanned by $\psi_0 \in R^n$; that is,

$R(f_x^0) = \{y \in R^n : \psi_0^T y = 0\}$.

If we write (1.1.1) as $f(x(s), \lambda(s)) = 0$, then differentiating with respect to s and evaluating at $s = s_0$; we obtain

$$f_x^0 \cdot x_0' + f_\lambda^0 \cdot \lambda_0' = 0. \qquad (1.\ 1.2)$$

where $f_\lambda^0 = f_\lambda(x_0, \lambda_0)$ and $'$ denotes differentiation with respect to s.

Differentiating again and evaluating at $s = s_0$, we obtain

$$f_x^0 \cdot x_0'' + f_\lambda^0 \cdot \lambda_0'' = -\left(f_{xx}^0 \cdot x_0' \cdot x_0' + 2f_{x\lambda}^0 \cdot x_0' \cdot \lambda_0' + f_{\lambda\lambda}^0 \cdot \lambda_0' \cdot \lambda_0'\right) \quad (1.\ 1.3)$$

where the subscript and superscript zero denotes evaluation at (x_0, λ_0) as before. Two cases will be considered depending on whether $\psi_0^T f_\lambda^0 = 0$ or not.

1.2 Turning Points

The first case to be considered is when $\psi_0^T f_\lambda^0 \neq 0$; that is, f_λ^0 is not in the range of f_x^0.

Let (x_0, λ_0) satisfy (1.1.1) and the assumptions H_1 and H_2 and $\psi_0^T f_\lambda^0 \neq 0$.

Now left multiplying (1.1.2) by ψ_0^T and since $\psi_0^T f_x^0 = 0$, this will imply that

$$\psi_0^T f_\lambda^0 \cdot \lambda_0' = 0,$$

thus

$$\lambda_0' = 0. \tag{1. 2.1}$$

Then equation (1.1.2) becomes $f_x^0 \cdot x_0' = 0$; that is, $x_0' \in N(f_x^0)$ or

$$x_0' = \alpha\phi_0. \tag{1. 2.2}$$

Now from (1.2.1), (1.2.2) and (1.1.3), we get

$$f_x^0 \cdot x_0'' + f_\lambda^0 \cdot \lambda_0'' = -f_{xx}^0 \cdot \phi_0 \cdot \phi_0. \tag{1. 2.3}$$

Applying ψ_0^T on equation (1.2.3) from left gives

$$\psi_0^T f_x^0 \cdot x_0'' + \psi_0^T f_\lambda^0 \cdot \lambda_0'' = -\psi_0^T f_{xx}^0 \cdot \phi_0 \cdot \phi_0. \tag{1. 2.4}$$

Since $\psi_0^T f_\lambda^0 \neq 0$ and $\psi_0^T f_x^0 = 0$, then equation (1.2.4) implies

$$\lambda_0'' = \frac{-\psi_0^T f_{xx}^0 \cdot \phi_0 \cdot \phi_0}{\psi_0^T f_\lambda^0}. \qquad (1.\ 2.5)$$

Definition 1.2.1 *(x_0, λ_0) is called a simple (quadratic) turning point if and only if, $\psi_0^T f_{xx}^0 \phi_0 \phi_0 \neq 0$. In other words, (x_0, λ_0) is simple turning point, if and only if, $\lambda_0' = 0$ and $\lambda_0'' \neq 0$.*

The term "simple" was used in Moore and Spence[14] and Spence and Werner[17] while the term "quadratic" was used in Jepson and Spence[9]. By differentiating (1.1.3) with respect to s and evaluating at $s = s_0$, we get

$$\begin{aligned} f_{xx}^0 \cdot x_0' \cdot x_0'' &+ f_{x\lambda}^0 \cdot \lambda_0' \cdot x_0'' + f_x^0 \cdot x_0''' + f_{\lambda\lambda}^0 \cdot \lambda_0' \cdot \lambda_0'' \\ &+ f_{\lambda x}^0 \cdot x_0' \lambda_0'' + f_\lambda^0 \cdot \lambda_0''' = -f_{xxx}^0 \cdot x_0' \cdot x_0' \cdot x_0' \\ &- f_{xx\lambda}^0 \cdot \lambda_0' \cdot x_0' \cdot x_0' - 2 f_{xx}^0 \cdot x_0' \cdot x_0'' - 2 f_{x\lambda\lambda}^0 \cdot \lambda_0' \cdot x_0' \cdot \lambda_0' \\ &- 2 f_{x\lambda x}^0 \cdot x_0' \cdot x_0' \cdot \lambda_0' - 2 f_{x\lambda}^0 \cdot x_0' \cdot \lambda_0'' - 2 f_{x\lambda}^0 \cdot x_0'' \cdot \lambda_0' \\ &- 2 f_{\lambda\lambda\lambda}^0 \cdot \lambda' \cdot \lambda' \cdot \lambda' - f_{\lambda\lambda x}^0 \cdot x_0' \cdot \lambda_0' \cdot \lambda_0' - 2 f_{\lambda\lambda}^0 \cdot \lambda_0' \cdot \lambda_0''. \end{aligned} \qquad (1.\ 2.6)$$

Since $\lambda_0' = 0$, all terms in (1.2.6) which contains λ_0' will vanish. Also, if $\psi_0^T f_{xx} \cdot \phi_0 \cdot \phi_0 = 0$, then $\lambda_0'' = 0$ and (1.2.6) becomes

$$f_{xx}^0 \cdot \phi_0 \cdot x_0'' + f_x^0 \cdot x_0''' + f_\lambda^0 \lambda_0''' \;=\; -f_{xxx}^0 \cdot \phi_0 \cdot \phi_0 \cdot \phi_0$$

$$-2f_{xx}^0 \cdot \phi_0 \cdot x_0''. \qquad (1.\ 2.7)$$

Applying ψ_0^T on equation (1.2.7) from left and since $\psi_0^T f_x^0 = 0$ and $\psi_0^T f_\lambda^0 \neq 0$, we obtain

$$\lambda_0''' = -(\psi_0^T f_{xxx}^0 \cdot \phi_0 \cdot \phi_0 \cdot \phi_0 + 3\psi_0^T f_{xx}^0 \cdot \phi_0 \cdot v_0)/\psi_0^T f_\lambda^0, \qquad (1.\ 2.8)$$

where $v_0 = x_0''$ and v_0 satisfies

$$f_x^0 \cdot v_0 = -f_{xx}^0 \cdot \phi_0 \cdot \phi_0, \qquad v_0 \in V \qquad (1.\ 2.9)$$

where V is the complement of $\{\phi_0\}$ in R^n, since $f_x^0 \cdot v_0 \neq 0$; that is $v_0 \notin N(f_x^0)$. Note that equation (1.2.9) can be obtained by setting $\lambda_0'' = 0$ in (1.2.3).

Definition 1.2.2 *A turning point (x_0, λ_0) is called cubic if and only if $\psi_0^T f_{xx}^0 \cdot \phi_0 \cdot \phi_0 = 0$ and $\psi_0^T f_{xxx}^0 \cdot \phi_0 \cdot \phi_0 \cdot \phi_0 + 3\psi_0^T f_{xx}^0 \phi_0 v_0 \neq 0$, where v_0 is as in equation (1.2.9). In other words, a turning point is called cubic if and only if $\lambda_0' = 0, \lambda_0'' = 0$ and $\lambda''' \neq 0$,*

1.3 Bifurcation Points

The second case to be considered is when $\psi_0^T f_\lambda^0 = 0$, that is f_λ^0 is in the range of f_x^0.

Definition 1.3.1 *(x_0, λ_0) is called a simple bifurcation point of (1.1.1) if $\psi_0^T f_\lambda^0 = 0$ and there exist two transversal solution curves intersecting at (x_0, λ_0).*

Now define $z_0 \in V$ by

$$f_x^0 \cdot z_0 = -f_\lambda^0. \tag{1. 3.1}$$

Using assumption H_1 together with (1.3.1) and (1.1.2) we will have

$$\lambda_0' = \beta_1 \tag{1. 3.2}$$

and

$$x_0' = \beta_0 \phi_0 + \beta_1 z_0. \tag{1. 3.3}$$

To explain that consider (1.1.2),+

$$\begin{aligned} f_x^0 x_0' + f_\lambda^0 \lambda_0' &= f_x^0(\beta_0 \phi_0 + \beta_1 z_0) + f_\lambda^0 \beta_1 \\ &= \beta_0 f_x^0 \phi_0 + \beta_1 f_x^0 z_0 + f_\lambda^0 \beta_1, \end{aligned}$$

$$
\begin{aligned}
&= \beta_1 f_x^0 z_0 + f_\lambda^0 \beta_1 \quad \text{since } \phi_0 \in N(f_x^0) \\
&= 0 \quad by (1.3.1).
\end{aligned}
$$

Now substituting (1.3.2) and (1.3.3) in (1.1.3), we obtain

$$
\begin{aligned}
f_x^0 \cdot x_0'' + f_\lambda^0 \cdot \lambda_0'' &= -\Big(f_{xx}^0 (\beta_0 \phi_0 + \beta_1 z_0)(\beta_0 \phi_0 + \beta_1 z_0) \\
&\quad + 2 f_{x\lambda}^0 (\beta_0 \phi_0 + \beta_1 z_0)\, \beta_1 + f_{\lambda\lambda}^0 \beta_1 \beta_1 \Big),
\end{aligned}
$$

expanding leads to

$$
\begin{aligned}
f_x^0 \cdot x_0'' + f_\lambda^0 \lambda_0'' &= -\beta_0^2 f_{xx}^0 \cdot \phi_0 \cdot \phi_0 - 2\beta_0 \beta_1 f_{xx}^0 \cdot \phi_0 \cdot z_0 \\
&\quad -\beta_1^2 f_{xx}^0 \cdot z_0 \cdot z_0 - 2\beta_0 \beta_1 f_{x\lambda}^0 \cdot \phi_0 \\
&\quad -2\beta_1^2 f_{x\lambda}^0 \cdot z_0 - f_{\lambda\lambda}^0 \beta_1^2. \qquad (1.\ 3.4)
\end{aligned}
$$

Left multiplying (1.3.4) by ψ_0^T we obtain

$$
\begin{aligned}
0 &= \beta_0^2 \psi_0^T f_{xx}^0 \cdot \phi_0 \cdot \phi_0 + 2\beta_0 \beta_1 \left(\psi_0^T f_{xx}^0 \cdot \phi_0 \cdot z_0 + \psi_0^T f_{x\lambda}^0 \cdot \phi_0 \right) \\
&\quad + \beta_1^2 \left(\psi_0^T f_{xx}^0 \cdot z_0 \cdot z_0 + 2\psi_0^T f_{x\lambda}^0 \cdot z_0 + \psi_0^T f_{\lambda\lambda}^0 \right). \qquad (1.\ 3.5)
\end{aligned}
$$

Now let

$$
a = \psi_0^T \cdot f_{xx}^0 \phi_0 \cdot \phi_0 \qquad (1.\ 3.6)
$$

$$b = \psi_0^T f_{xx}^0 \cdot \phi_0 \cdot z_0 + \psi_0^T f_{x\lambda}^0 \phi_0 \quad (1.\ 3.7)$$

$$c = \psi_0^T f_{xx}^0 \cdot z_0 \cdot z_0 + 2\psi_0^T f_{x\lambda}^0 \cdot z_0 + \psi_0^T f_{\lambda\lambda}^0. \quad (1.\ 3.8)$$

Then equation (1.3.5) has the form

$$a\beta_0^2 + 2b\beta_0\beta_1 + c\beta_1^2 = 0, \quad (1.\ 3.9)$$

the quadratic formula gives two roots β_0, namely,

$$\begin{aligned}\beta_0 &= \frac{-2b\beta_1 \mp \sqrt{4b^2\beta_1^2 - 4ac\beta_1^2}}{2a} \\ &= \frac{-2b\beta_1 \mp 2\beta_1\sqrt{b^2 - ac}}{2a}. \end{aligned} \quad (1.\ 3.10)$$

Defining $\Delta = b^2 - ac$, we will have three cases.

(i) If $\Delta > 0$ then equation (1.3.9) will have two distinct roots, and there exist two nontangential solution curves intersecting at (x_0, λ_0). This point will be called a simple bifurcation point. See Jepson and Spence[9] and Keller[11].

Definition 1.3.2 *A simple bifurcation point is called transcritical point if and only if* $a = \psi_0^T f_{xx}^0 \cdot \phi_0 \cdot \phi_0 \neq 0$.

Definition 1.3.3 *A simple bifurcation point is called a Pitch fork bifurcation point if and only if* $a = \psi_0^T f_{xx}^0 \cdot \phi_0 \cdot \phi_0 = 0$ and $\lambda_0' = 0$.

(ii) If $\Delta < 0$, then there are no real roots of (1.3.9), and (x_0, λ_0) is an isolated conjugate point.

(iii) If $\Delta = 0$, with at least one of a, b or being nonzero, then (1.3.9) has one nontrivial root. This point is called a cusp point, and the solution arcs have second order contact at (x_0, λ_0), the slope at the singular point of higher-order contact is given by (1.3.10). See Jepson and Spence[9].

Chapter 2

CHARACTERIZATION OF SIMPLE SINGULARITIES

To solve the system (1.1.1) for any given λ different from the singular point, one may be able to characterize the full solution set

$$M = \{x \in R^n; \quad f(x, \lambda) = 0\}.$$

Newton's method can be used to solve this sytem. However, for singular points, that is, a point (x_0, λ_0) that makes $f_x(x, \lambda)$ singular, Newton's method will fail. For this case we define S, the set of singular points of f, to be those points where the rank of f_x^0 drops by one.

The intersection of M and S is what we are after which will be isolated points.

2.1 Characterization of Simple (Quadratic) Turning Points

We will assume that (x_0, λ_0) is a simple (quadratic) turning point of the system (1.1.1) and the rank of (f_x^0) drops by one while the full Jacobian maintains rank n. This is due to assumptions (H_1) and (H_2) in section (1.1), and the fact that f_λ^0 is not in the range of (f_x^0), in the simple turning point case.

The main problem will be the characterization of the singular manifold S and then numerically one can compute the intersection between M and S, where M and S are as defined above. The set M forms a smooth $p = n + 1 - n = 1$ dimensional manifold, while S will form a smooth $n+1-p$ dimensional manifold. This means that the dimensions of M and S add up to $n+1$ and the intersection of M and S will be isolated points. Now if an equation or equations are given that characterize S, then the system (1.1.1) will be augmented by that equation or equations.

The approach we will use to characterize S in the one used in Attili[2,3] and Greiwank and Reddien[7]. Such approach produces an augmented system for which the singular point is an isolated solution since it was proved there that the Jacobian of the system is nonsingular at the simple (quadratic) turning point. A general framework for generating such equation is given in the following Lemma which is a special case of Lemma given in Griewank and Reddien[7] and Attili[1,3].

Lemma 2.1.1 *Let T and R be continuously differentiable vector functions mapping $D \subset R^n \times R$ into $R^n \times R$ and R^n respectively so that the square matrix*

$$\begin{bmatrix} f_x & R \\ T^T & 0 \end{bmatrix} \tag{2. 1.1}$$

is nonsingular for all (x, λ) in D. Then there are unique vector functions u, v and g such that

(i)

$$f_x \cdot v = -R \cdot g, \qquad T^T v = 1$$

(ii)

$$u^T \cdot f_x = -g \cdot T^T, \qquad u^T \cdot R = 1 \tag{2. 1.2}$$

where v is $n \times 1$, g is 1×1 (scalar function) and u is $n \times 1$. It follows then that

$$\text{rank} \quad (f_x) = n - 1 \tag{2. 1.3}$$

if and only if $g(x, \lambda) = 0$, moreover u, v and g are continuously differentiable, and in particular

$$g' = (-u^T f_x v)' = -u^T f_x' v - u^T R' g^T - g^T (T')^T v \tag{2. 1.4}$$

where the prime denotes differentiation in x or λ.

Proof:

Consider

$$\text{(a)} \qquad \begin{bmatrix} f_x & R \\ T^T & 0 \end{bmatrix} \begin{bmatrix} v \\ g \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$\text{(b)} \qquad \begin{bmatrix} u^T g \end{bmatrix} \begin{bmatrix} f_x & R \\ T^T & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \end{bmatrix}. \qquad (2.\ 1.5)$$

Since the matrix on the left hand side of (2.1.5) is nonsingular, then the existence, uniqueness and continuous differentiability of v, u and g follow directly.

From (2.1.5-a) we obtain (2.1.2-i), and from (2.1.5-b) we obtain (2.1.2-ii).

Left multiplying (2.1.2-i) by u^T and right multiplying (2.1.2-ii) by v, and recalling $u^T R = 1$ and $T^T v = 1$ we find

$$g = -u^T f_x v. \qquad (2.\ 1.6)$$

Differentiating (2.1.6), we obtain

$$g' = -\left[(u')^T \cdot f_x \cdot v + u^T \cdot f_x' \cdot v + u^T \cdot f_x \cdot v'\right]. \qquad (2.\ 1.7)$$

Differentiating (2.1.2-i) and left multiplying by u^T, we obtain

$$u^T \cdot f_x' v + u^T \cdot f_x v' = -u^T \cdot R' g - u^T \cdot R g'. \qquad (2.\ 1.8)$$

Differentiating (2.1.2-ii) and right multiplying by v, we obtain

$$(u')^T \cdot f_x \cdot v + u^T f_x' \cdot v = -g' T^T \cdot v - g(T')^T \cdot v. \qquad (2.\ 1.9)$$

Adding (2.1.8) to (2.1.9) and recalling $u^T R = 1$ and $T^T v = 1$, we obtain

$$u^T \cdot f_x' \cdot v' + (u')^T \cdot f_x \cdot v = -2u^T \cdot f_x' v - 2g' - u^T \cdot R' g - g(T')^T. \quad (2.\ 1.10)$$

Substituting (2.1.10) in (2.1.7), we obtain (2.1.4); that is

$$g' = -u^T \cdot f_x' v - u^T \cdot R' \cdot g - g(T')^T v.$$

If R and T were chosen to be constants or (x_0, λ_0) is the singular point, then equation (2.1.4) will reduce to $g' = -u^T \cdot f_x' \cdot v$, also notice that $u = \psi_0$, where span $\{\psi_0\} =$ null $\left\{f_x^{0^T}\right\}$ and $v = \phi_0$ where span$\{\phi_0\} =$ null $\{f_x^0\}$ at the singular point.

Equation (2.1.3) of Lemma (2.1.1) shows that the singular set S is indeed the zero set for g. So finding the simple turning point will consist of augmenting system (1.1.1) with $g = 0$ to obtain

$$f(x,\lambda) = 0$$

$$g(x,\lambda) = 0. \qquad (2.\ 1.11)$$

To solve this system, Newton's method can be used. For that reason one would like to guarantee the existence of an isolated solution of (2.1.11). Looking at the Jacobian of the system (2.1.11); that is,

$$\begin{bmatrix} f_x & f_\lambda \\ g_x & g_\lambda \end{bmatrix},$$

and considering

$$\begin{bmatrix} f_x^0 & f_\lambda^0 \\ g_x^0 & g_\lambda^0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This implies that

$$\text{(i)} \qquad f_x \cdot \hat{x} + \alpha f_\lambda^0 = 0$$

$$\text{(ii)} \qquad g_x^0 \cdot \hat{x} + \alpha g_\lambda^0 = 0. \qquad (2.\ 1.12)$$

Since f_λ^0 is not in the range of (f_x^0) and left multiplying (2.1.12-i) by ψ_0^T we will have $\alpha = 0$ which implies $\hat{x} \in N(f_x^0)$; that is, $\hat{x} = \beta\phi_0$. As a result (2.1.12-ii) becomes $\beta g_x^0\phi_0 = 0$, but we know that $g_x^0 = -\psi_0^T f_{xx}^0\phi_0$, which means that $\beta = 0$ if and only if the condition $\psi_0^T f_{xx}^0\phi_0\phi_0 \neq 0$ is satisfied.

This shows that such condition is essential to make the Jacobian of the determining system (2.1.11) nonsingular. From this discussion we will have the following theorem.

Theorem 2.1.1 *If (x_0, λ_0) is a quadratic turning point of $f(x, \lambda) = 0$, then (x_0, λ_0) is strongly isolated solution to (2.1.11) in the sense that the Jacobian of the determining system (2.1.11) is nonsingular.*

The proof is straight forward as discussed above and for more details on the proof see Attili[1] and the references there.

2.2 Characterization of Simple Bifurcation Points

In the simple turning point case, and in order to numerically compute the singular point, the under determined system $f(x, \lambda) = 0$ was augmented with an equation $g(x, \lambda) = 0$, that enforces the singularity of f_x^0. Here we will also use the same approach with slight modification to compute the simple bifurcation point.

To explain how to characterize the simple bifurcation point in order to numerically be able to compute it, let us recall that as we mentioned in the previous chatper, f^0_λ is in the range of f^0_x and again f^0_x has a one-dimensional null space. With this in mind, we will have the following lemma.

Lemma 2.2.1 *Let* $f^0_{\bar{x}} = (f^0_x, f^0_\lambda), \bar{x} = (x, \lambda)$. *Then* Null $(f^0_{\bar{x}}) =$ span $\{\Phi_0, \Phi_1\}$ *where* $\Phi_0 = (\phi_0, 0)^T, \Phi_1 = (\phi_1, \gamma)^T$ and ϕ_1 *is such that*

$$f^0_x \cdot \phi_1 + \gamma f^0_\lambda = 0, \;\; \phi_1 \in \text{ Range } (f^0_x) \text{ and } \gamma \in R.$$

Proof:

Consider

$$\left(f^0_x, f^0_\lambda\right) \cdot (\beta\Phi_0, \alpha\Phi_1) = \left(\left(f^0_x, f^0_\lambda\right) \cdot \begin{pmatrix} \beta\phi_0 \\ 0 \end{pmatrix}, \left(f^0_x, f^0_\lambda\right) \cdot \begin{pmatrix} \alpha\phi_1 \\ \alpha\gamma \end{pmatrix}\right),$$

which in expanded from equals

$$= \left(\beta f^0_x \phi_0, \quad \alpha f^0_x \phi_1 + \alpha\gamma f^0_\lambda\right) = (0, 0),$$

the first part and second part vanish since $\phi_0 \in$ Null (f^0_x) and $f^0_x \phi_1 + \gamma f^0_\lambda = 0$ respectively by the statement of the Lemma. Hence Null $(f^0_{\bar{x}}) =$ span$\{\Phi_0, \Phi_1\}$.

It follows from Lemma (2.2.1) that the solution manifold $f^{-1}(0)$ is a p-dimensional manifold ($p = 2$) and the singular manifold an $n + 1 - p$ dimensional manifold. See Attili[4].

For numerical calculations the singularity will be unfolded as $f(x,\lambda)+\alpha r=0$, where r is not in the range of (f_x, f_λ), $\alpha \in R$ is a perturbation parameter. Also since $p=2$, then $g(x,\lambda)$ which is as in Lemma (2.1.1) will have two components g_1 and g_2. To determine these two components, we solve the systems

$$\begin{pmatrix} f_{\bar{x}} & R \\ T^T & 0 \end{pmatrix}\begin{pmatrix} v \\ g \end{pmatrix}=\begin{pmatrix} f_x & f_\lambda & R \\ T^T & & 0 \end{pmatrix}\cdot\begin{pmatrix} v_1 & v_2 \\ g_1 & g_2 \end{pmatrix}=\begin{pmatrix} 0 \\ I_2 \end{pmatrix} \qquad (2.\ 2.1)$$

and

$$\left(u^T, g\right)\begin{pmatrix} f_x & f_\lambda & R \\ T^T & & 0 \end{pmatrix}=(0,1), \qquad (2.\ 2.2)$$

here T^T will be a $2\times(n+1)$ matrix, u^T is a $1\times(n+1)$ vector, $g=(g_1,g_2)$ where g_1 and g_2 are scalar quantities, R will be an $(n\times 1)$ vector and I_2 is the 2×2 identity matrix.

As a result the two components of g will be

$$g_1 = -u^T f_x v_1$$

$$g_2 = -u^T f_x v_2. \qquad (2.\ 2.3)$$

The above two components are obtained exactly as outlined in the proof of Lemma (2.1.1).

The extended system for the calculation of simple bifurcation point will be

$$f(x,\lambda)+\alpha r \;=\; 0$$

$$
\begin{aligned}
g_1(x,\lambda) &= 0 \\
g_2(x,\lambda) &= 0,
\end{aligned}
\tag{2. 2.4}
$$

which is an $(n+2)$ system of equations in $(n+2)$ unknowns, namely (x, λ, α), where α will be identically zero at the simple bifurcation point. Note that the two components of g are independent of the perturbation parameter α, hence the Jacobian of the determining system (2.2.4) will be

$$
\begin{pmatrix} f_x & f_\lambda & r \\ g_{1x} & g_{1\lambda} & 0 \\ g_{2x} & g_{2\lambda} & 0 \end{pmatrix}. \tag{2. 2.5}
$$

Now considering

$$
\begin{pmatrix} f_x & f_\lambda & r \\ g_{1x} & g_{1\lambda} & 0 \\ g_{2x} & g_{2\lambda} & 0 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \alpha \\ \beta_1 \end{pmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.
$$

This implies that

(i)

$$
f_x \cdot \hat{x} + \alpha f_\lambda + \beta_1 r = 0
$$

(ii)

$$
g_{1x} \cdot \hat{x} + g_{1\lambda}\alpha = 0
$$

(iii)

$$
g_{2x} \cdot \hat{x} + g_{2\lambda}\alpha = 0, \tag{2. 2.6}
$$

since r is not in the range of (f_x, f_λ) we will have $\beta_1 = 0$. As deduced earlier in this section and at the simple bifurcation point the two components of g are

$$g_i = -\psi_0^T f_x^0 \phi_{i-1}; \quad i = 1, 2. \tag{2. 2.7}$$

Differentiating and if T and R were chosen to be constants as was discussed in section (2.1), we obtain

$$g_i' = \psi_0^T f_x^{0'} \phi_{i-1}; \quad i = 1, 2 \tag{2. 2.8}$$

where the prime denotes differentiation with respect to x or λ. Substituting (2.2.7) and (2.2.8) in (2.2.6), we will have the following result.

Theorem 2.2.1 *The simple bifurcation point (x_0, λ_0) is an isolated solution of (2.2.4) if and only if the determinant of the 2×2 symmetric matrix*

$$\begin{pmatrix} \psi_0^T f_{xx}^0 \phi_0 \phi_0 & \psi_0^T f_{xx}^0 \phi_0 \phi_1 \\ \psi_0^T f_{xx}^0 \phi_0 \phi_1 & \psi_0^T f_{xx}^0 \phi_1 \phi_1 \end{pmatrix} \tag{2. 2.9}$$

is nonzero where ϕ_0 spans null (f_x^0) and ϕ_1 as defined in Lemma (2.2.1).

For more details on condition (2.2.9) see Keller[11] and Rheinboldt[15] and for more on the proof, see Griewank and Reddien[7].

2.3 Characterization of Cubic Turning Points

The problem under consideration is a two parameter problem of the form $f(x, \lambda, \mu) = 0$. With $\mu = \mu_0$ fixed, $f(x, \lambda, \mu_0) = 0$ will have a simple turning point with respect to λ. To compute such point one should extend the system $f(x, \lambda, \mu_0) = 0$ to obtain

$$F(y, \mu_0) = \begin{pmatrix} f(x, \lambda, \mu_0) \\ g(x, \lambda, \mu_0) \end{pmatrix}, \tag{2. 3.1}$$

where $y = (x, \lambda)$ and $g = -\psi_0^T \cdot f_x \phi_0$ as defined earlier. By theorem (2.1.2) the Jacobian of system (2.3.1) will be nonsingular, and so such simple turning point can be computed without any difficulty. The following theorem concludes that the numerical computation of the cubic turning points of $f(x, \lambda, \mu) = 0$ with respect to λ and μ fixed is simply calculating a simple turning point of the extended system in (2.3.1) with respect to μ.

Theorem 2.3.1 *Let $\psi_0^T f_{xxx}^0 \phi_1 \phi_0 = 0$ and $\psi_0^T (3 f_{xxx}^0 \phi_0 v_0 + f_{xxx}^0 \phi_0 \phi_0 \phi_0) \neq 0$ where $f_x^0 v_0 = -f_{xx}^0 \phi_0 \cdot \phi_0$. Let $\beta_0 f_\lambda^0 = -\psi_0^T f_{x\lambda}^0 \phi_0$, $\beta_0 f_x^0 = -\psi_0^T f_{xx}^0 \phi_0$ and assume $\beta_0 f_x^0 = -\beta_0 f_\mu^0 + \psi_0^T \cdot f_{x\mu}^0 \cdot \phi_0 \neq 0$. Then a cubic (double) turning point (x_0, λ_0, μ_0) of $f(x, \lambda, \mu) = 0$ with respect to λ and $\mu = \mu_0$ fixed corresponds to simple turning point (x_0, λ_0, μ_0) of $F(y, \mu) = 0$ with respect to μ.*

For the proof one can refer to Attili[1] and the references there. A similar result with its proof was stated in Spence and Werner[17] for a different but larger extension than the one in (2.3.1). Also more on the terminology can be found elsewhere, for example Jepson and Spence[9].

Thus one has to extend (2.3.1) again to obtain $\tilde{g}(x, \lambda, \mu)$ to correspond to $F(y, \mu)$ in the usual way explained earlier in Lemma (2.1.1), to obtain the system

$$\tilde{F}(y,\mu) = \begin{cases} f(x,\lambda,\mu) = 0 \\ g(x,\lambda,\mu) = 0 \\ \tilde{g}(x,\lambda,\mu) = 0. \end{cases} \tag{2. 3.2}$$

To determine $\tilde{g}$, we solve the system

$$\begin{pmatrix} f_x & f_\lambda & \tilde{R} \\ g_x & g_\lambda & \\ \tilde{T}^T & & 0 \end{pmatrix} \cdot \begin{pmatrix} \tilde{v} \\ \tilde{g} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{2. 3.3}$$

where $\tilde{T}^T$ will be a row vector, $\tilde{R}$ will be a colum vector, $\tilde{v}$ will be a column vector and $\tilde{g}$ a scalar function. The Jacobian of system (2.3.2) is

$$\begin{pmatrix} f_x & f_\lambda & f_\mu \\ g_x & g_\lambda & g_\mu \\ \tilde{g}_x & \tilde{g}_\lambda & \tilde{g}_\mu \end{pmatrix} \tag{2. 3.4}$$

Chapter 3

NUMERICAL DETAILS

The problem under consideration is a two point boundary value problem of the form

$$\bar{F}(z,\lambda) = \begin{cases} z''(x) + f(x, z(x), \lambda) = 0 & a \le x \le b \\ z(a) = \alpha, \quad z(b) = \beta \end{cases}$$

where $\bar{F}(z,\lambda)$ is a nonlinear operator from $R^n \times R$ into R^n and λ is a real parameter. We will use finite differences as detailed in the following section efficiently. We will make use of the special structure of the resulting systems through block elimination. Many authors used different methods to discretize the above problem to obtain a finite dimentional one. For example Attili[1] has used multiple shooting, Kikuchi[12] used finite elements and Weiss[19] used difference approximations. It is well known that finite differences have a better stability chracteristics, but they generally require more work to obtain a specified accuracy. In this chapter we will discuss how to overcome this particular problem which will be done using block-elimination together with Richardson extrapolation.

3.1 Discretization by Finite - Differences

To change the above infinite-dimensional problem into a finite-dimensional one, we use a finite-difference method to discretize it.

We start by dividing the interval $[a, b]$ into $(N+1)$ subintervals whose endpoints are at $x_i = a + ih$ for $i = 1, 2, ..., N$, where $h = (b-a)/(N+1)$, and replace $z''(x_i)$ in each of the equations

$$z''(x_i) + f(x_i, z(x_i), \lambda) = 0,$$

by the appropriate centered-difference formula given by

$$z''(x_i) = \frac{1}{h^2}\left[z(x_{i+1}) - 2z(x_i) + z(x_{i-1})\right] - \frac{h^2}{12} z^{(4)}(\xi_i) \qquad (3.\ 1.1)$$

where $\xi_i \in (x_{i-1}, x_{i+1})$, to obtain

$$\frac{z(x_{i+1}) - 2z(x_i) + z(x_{i-1})}{h^2} + f\left(x_i, z(x_i), \lambda\right) - \frac{h^2}{12} z^{(4)}(\xi) = 0 \qquad (3.\ 1.2)$$

for each $i = 1, 2, \cdots, N$ and again for some ξ in the interval (x_{i-1}, x_{i+1}).

A finite-difference method reults when the error terms are deleted and using the boundary conditions $z(a) = \alpha$ and $z(b) = \beta$ to define

$$z(x_0) = \alpha \qquad z(x_{N+1}) = \beta,$$

and

$$\frac{z_{i+1} - 2z_i + z_{i-1}}{h^2} + f(x_i, z_i, \lambda) = 0,$$

for each $i = 1, 2, \cdots, N$, where z_i is the approximation of $z(x_i)$.

The $N \times N$ nonlinear system obtained from this method is given by

$$F(z,\lambda) = \begin{cases} \dfrac{-2z_1 + z_2 + \alpha}{h^2} + f(x_1, z_1, \lambda) & = 0, \\ \dfrac{z_1 - 2z_2 + z_3}{h^2} + f(x_2, z_2, \lambda) & = 0, \\ & \vdots \\ \dfrac{z_{N-2} - 2z_{N-1} + z_N}{h^2} + f(x_{N-1}, z_{N-1}, \lambda) & = 0, \\ \dfrac{z_{N-1} - 2z_N + \beta}{h^2} + f(x_N, z_N, \lambda) & = 0, \end{cases} \tag{3. 1.3}$$

which has the Jacobian matrix J given by

$$F_z(z,\lambda) = \begin{bmatrix} \frac{-2}{h^2}+a_1 & \frac{1}{h^2} & 0 & \cdots & \cdots & & 0 \\ \frac{1}{h^2} & \frac{-2}{h^2}+a_2 & \frac{1}{h^2} & & & & 0 \\ 0 & & & & & & \vdots \\ \vdots & & & & & & 0 \\ \vdots & & & & & & \frac{1}{h^2} \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{1}{h^2} & \frac{-2}{h^2}+a_N \end{bmatrix} \tag{3. 1.4}$$

where $a_i = f_{z_i}(x_i, z_i, \lambda)$ for each $i = 1, 2, \cdots, N$.

To compute a simple turning point we need to solve the system (2.1.11); that is,

$$\begin{aligned} F(z,\lambda) &= 0 \\ g(z,\lambda) &= 0, \end{aligned} \tag{3. 1.5}$$

where $F(z,\lambda)$ is the system given in (3.1.3). From Lemma (2.1.1), the determining system for g will be

$$
\begin{bmatrix}
\frac{-2}{h^2}+a_1 & \frac{1}{h^2} & 0 & \cdots & \cdots & & 0 & R_1 \\
\frac{1}{h^2} & \frac{-2}{h^2}+a_2 & \frac{1}{h^2} & & & & 0 & \vdots \\
0 & & & & & & & \vdots \\
\vdots & & & & & & 0 & \vdots \\
\vdots & & & & & & \frac{1}{h^2} & \\
0 & \cdots & \cdots & \cdots & 0 & \frac{1}{h^2} & \frac{-2}{h^2}+a_N & R_N \\
T_1 & T_2 & \cdots & \cdots & & & T_N & 0
\end{bmatrix}
\cdot
\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \\ g \end{bmatrix}
=
\begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix}
\tag{3. 1.6}
$$

where T and R are to be chosen as we described in Lemma (2.1.1), more details on the choice of T and R will be given at a later stage in this chapter. In a similar way, the determining system for u will be

$$
(u^T, g)\begin{bmatrix} J & R \\ T^T & 0 \end{bmatrix} = [0, 1] \tag{3. 1.7}
$$

If (z_0, λ_0) is a simple turning point of $F(z, \lambda) = 0$, then (z_0, λ_0) is a strongly isolated solution to (3.1.5), which is guaranteed by theorem (2.1.3); that is, the Jacobian matrix J_t of system (3.1.5) given by

$$
\begin{bmatrix}
\frac{-2}{h^2}+a_1 & \frac{1}{h_2} & 0 & \cdots & \cdots & & 0 & F_{1\lambda} \\
\frac{1}{h^2} & \frac{-2}{h^2}+a^2 & \frac{1}{h^2} & & & & 0 & \vdots \\
0 & & & & & & & \vdots \\
\vdots & & & & & & 0 & \vdots \\
\vdots & & & & & & \frac{1}{h^2} & \\
0 & \cdots & & \cdots\ \cdots & 0 & \frac{1}{h^2} & \frac{-2}{h^2}+a_N & F_{N\lambda} \\
g_{z1} & & & & & & g_{zN} & g_\lambda
\end{bmatrix},
\tag{3. 1.8}
$$

is nonsingular where $F_{i\lambda} = f_\lambda(x_i, z_i, \lambda)$. To compute the solution of (3.1.5), we will use Newton's method for nonlinear systems; that is, the linear systems to be solved are of the form

$$J_t Z = -F,$$

where

$$
\begin{aligned}
F \ = \ & \left(\frac{-2z_1 + z_2 + \alpha}{h^2} + f(x_1, z_1, \lambda), \cdots, \right. \\
& \left. \frac{z_{N-1} - 2z_N + \beta}{h^2} + f(x_N, z_N, \lambda), g\right).
\end{aligned}
$$

In the bifurcation point case, and to numerically calculate the singular point, we will solve the system

$$\begin{aligned} F(z,\lambda)+\gamma r &= 0 \\ g_1(z,\lambda) &= 0 \\ g_2(z,\lambda) &= 0, \end{aligned} \qquad (3.\ 1.9)$$

where γ is the perturbation parameter which will be identically zero at the singular point. The perturbed form of $F(z,\lambda)$; that is, the first part of (3.1.9) is what people usually use to unfold the singularity.

To determine g_1 and g_2, we solve the following systems

$$\begin{bmatrix} \frac{-2}{h^2}+a_1 & \frac{1}{h^2} & 0 & & & 0 & F_{1\lambda} & R_1 \\ \frac{1}{h^2} & \frac{-2}{h^2}+a_2 & \frac{1}{h^2} & & & 0 & \vdots & \vdots \\ 0 & & & & & & & \\ \vdots & \vdots & & & & 0 & \vdots & \\ \vdots & \vdots & & & & \frac{1}{h^2} & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & \frac{1}{h^2} & \frac{-2}{h^2}+a_N & F_{N\lambda} & R_N \\ t_{11} & \cdots & \cdots & & & & t_{1N+1} & 0 \\ t_{21} & \cdots & \cdots & & & & t_{2N+1} & 0 \end{bmatrix} \cdot \begin{bmatrix} v_1 & w_1 \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ v_{N+1} & w_{N+1} \\ g_1 & g_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (3.1.10)$$

Here v, w spans $N(f_z, f_\lambda)$.

Similarly to calculate u, we solve

$$(u^T, g)\begin{bmatrix} J & F_\lambda & R \\ & T^T & 0 \end{bmatrix} = [0,1] \tag{3. 1.11}$$

where R is not in the range of (F_z, F_λ), T^T will be a $2 \times (N+1)$ matrix and $g = (g_1, g_2)$.

The bifurcation point (z_0, λ_0) will be an isolated solution of the system (3.1.9) as was stated in theorem (2.2.2); that is, the Jacobian of the system in (3.1.9)

$$J_b = \begin{bmatrix} \frac{-2}{h^2}+a_1 & \frac{1}{h^2} & 0 & \cdots & \cdots & \cdots & 0 & f_{1\lambda} & r_1 \\ \frac{1}{h^2} & \frac{-2}{h^2}+a_2 & \frac{1}{h^2} & & & & 0 & \vdots & \vdots \\ 0 & & & & & & & \vdots & \vdots \\ \vdots & & & & & & 0 & \vdots & \vdots \\ \vdots & & & & & & \frac{1}{h^2} & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{1}{h^2} & \frac{-2}{h^2}+a_N & F_{N\lambda} & r_N \\ g_{1z_1} & \cdots & \cdots & \cdots & & & g_1 z_N & g_{1\lambda} & 0 \\ g_{2z_1} & \cdots & \cdots & \cdots & & & g_2 z_N & g_{2\lambda} & 0 \end{bmatrix} \tag{3. 1.12}$$

is nonsingular. Again Newton's method for nonlinear systems will be used to solve (3.1.9) which requires at each iteration the solution of the $(N+2)\times(N+2)$ linear system

$$J_b Z = -F \tag{3. 1.13}$$

where

$$\begin{aligned} F &= \Big(\frac{-2z_1+z_2+\alpha}{h^2}+f(x_1,z_1,\lambda),\cdots, \\ &\quad \frac{z_{N-1}-2z_N+\beta}{h^2}+f(x_N,z_N,\lambda),g_1,g_2\Big). \end{aligned}$$

Finally to compute cubic turning points, we need to solve the system in (2.3.2); that is,

$$\tilde{F}(y,\mu)=\begin{cases} F(z,\lambda,\mu) &= 0 \\ g(z,\lambda,\mu) &= 0 \\ \tilde{g}(z,\lambda,\mu) &= 0. \end{cases} \tag{3. 1.14}$$

From Lemma (2.1.1) the determining systems for g and $\tilde{g}$ will be respectively

$$
\begin{bmatrix}
\frac{-2}{h^2}+a_1 & \frac{1}{h^2} & 0 & \cdots & \cdots & 0 & R_1 \\
\frac{1}{h^2} & \frac{-2}{h^2}+a_2 & \frac{1}{h^2} & & & 0 & \vdots \\
0 & & & & & & \\
\vdots & & & & & 0 & \vdots \\
\vdots & & & & & \frac{1}{h^2} & \\
0 & & & 0 & \frac{1}{h^2} & \frac{-2}{h^2}+a_N & R_N \\
t_1^T & \cdots & \cdots & & & T_N^T & 0
\end{bmatrix}
\cdot
\begin{bmatrix} v_1 \\ \vdots \\ \vdots \\ \vdots \\ v_N \\ g \end{bmatrix}
=
\begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix}
\tag{3.1.15}
$$

and

$$
\begin{bmatrix}
\frac{-2}{h^2}+a_1 & \frac{1}{h^2} & 0 & \cdots & & 0 & F_{1\lambda} & \tilde{R}_1 \\
\frac{1}{h^2} & \frac{-2}{h^2}+a_2 & \frac{1}{h^2} & & & 0 & \vdots & \vdots \\
0 & & & & & & \vdots & \vdots \\
\vdots & & & & & 0 & & \\
\vdots & & & & & \frac{1}{h^2} & \vdots & \vdots \\
0 & & & 0 & \frac{1}{h^2} & \frac{-2}{h^2}+a_N & F_{N\lambda} & \tilde{R}_N \\
g_{z1} & \cdots & \cdots & & & g_{zN} & g_\lambda & \tilde{R}_{N+1} \\
\tilde{t}_1 & \cdots & \cdots & & & & \tilde{t}_{N+1} & 0
\end{bmatrix}
\cdot
\begin{bmatrix} \tilde{v} \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \tilde{v}_N \\ \tilde{v}_{N+1} \\ \tilde{g} \end{bmatrix}
=
\begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix}
\tag{3. 1.16}
$$

Also $T, \tilde{T}, R$ and $\tilde{R}$ are to be chosen as was described in Lemma (2.1.1) and more details will be given at a later stage in this chapter. To find u and $\tilde{u}$, the systems to be solved will be respectively

$$(u^T, g)\begin{bmatrix} J & R \\ T^T & 0 \end{bmatrix} = [0, 1] \tag{3. 1.17}$$

and

$$(\tilde{u}^T, \tilde{g})\begin{bmatrix} J & F_\lambda & \\ & & \tilde{R} \\ g_z & g_\lambda & \\ & \tilde{T}^T & 0 \end{bmatrix} = [0, 1] \tag{3. 1.18}$$

The Jacobian of the system (3.1.14) is given by

$$\begin{bmatrix} \frac{-2}{h^2}+a_1 & \frac{1}{h^2} & 0 & \cdots & \cdots & & 0 & F_{1\lambda} & F_{1\mu} \\ \frac{1}{h^2} & \frac{-2}{h^2}+a_2 & \frac{1}{h^2} & & 0 & & \vdots & \vdots & \vdots \\ 0 & & & & & & \vdots & \vdots & \vdots \\ \vdots & & & & & & 0 & \vdots & \vdots \\ \vdots & & & & & & \frac{1}{h^2} & & \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{1}{h^2} & \frac{-2}{h^2}+a_N & F_{N\lambda} & F_{N\mu} \\ g_{z1} & \cdots & \cdots & \cdots & & & g_{zN} & g_\lambda & g_\mu \\ \tilde{g}_{z1} & \cdots & \cdots & \cdots & & & \tilde{g}_{zN} & \tilde{g}_\lambda & \tilde{g}_\mu \end{bmatrix}, \tag{3. 1.19}$$

is nonsingular at the cubic turning point (z_0, λ_0, μ_0), see the result of Theorem (2.3.1). In this case the system to be solved which is similar to (3.1.13) has the form

$$J_c Z = -F, \tag{3. 1.20}$$

where

$$\begin{aligned} F \;=\; & \Big(\frac{-2z_1 + z_2 + \alpha}{h^2} + f(x_1, z_1, \lambda), \cdots, \\ & \frac{z_{N-1} - 2z_N + \beta}{h^2} + f(x_N, z_N, \lambda), g, \bar{g}\Big) . \end{aligned}$$

To apply the extrapolation technique, again consider the system

$$\begin{cases} z''(x) + f(x, z(x), \lambda) = 0 & a \leq x \leq b \\ z(a) = \alpha, & z(b) = \beta. \end{cases} \tag{3. 1.21}$$

Suppose Z is an isolated solution of this system and $Z_e(h)$ is the formula that produces approximate solution of order $O(h^2)$ to Z; that is, the finite-difference formulas. Assume that the error form for the approximation of $Z_e(h)$ to Z can be expressed as

$$Z = Z_e(h) + k_1 h^2 + O(h^4), \tag{3. 1.22}$$

where k_1 does not depend on h. To get more accurate approximations, we will follow the Richardson approximation, see Keller [9]. It starts by finding a new solution Z_e using the finite-difference approximation with a step size of length $(h/2)$ which is half the previous step to obtain

$$Z = Z_e(h/2) + k_1 \frac{h^2}{4} + O(h/2)^4. \qquad (3.\ 1.23)$$

Multiplying equation (3.1.23) by 4 and subtracting equation (3.1.22) from it gives

$$Z = \frac{4Z_e(h/2) - Z_e(h)}{3} + O(h^4), \qquad (3.\ 1.24)$$

define

$$Z_{e1} = Z_e(h) \qquad \text{and}$$

$$Z_{e2}(h) = \frac{4Z_{e1}(h/2) - Z_{e1}(h)}{3}, \qquad (3.\ 1.25)$$

then equation (3.1.24) becomes

$$Z = Z_{e2}(h) + O(h^4), \qquad (3.\ 1.26)$$

where $Z_{e2}(h)$ has the higher order $O(h^4)$.

If also a number k_2 independent of h exists, so that equation (3.1.22) can be expressed as

$$Z = Z_e(h) + k_1h^2 + k_2h^4 + O(h^6), \qquad (3.\ 1.27)$$

then equation (3.1.26) becomes

$$Z = Z_{e2}(h) - \frac{1}{4}k_2h^4 + O(h^6). \qquad (3.\ 1.28)$$

Replacing h with $h/2$ again in (3.1.28) gives

$$Z = Z_{e2}(h/2) - \frac{1}{64}k_2h^4 + O(h^6). \qquad (3.\ 1.29)$$

Multiplying (3.1.29) by 16 and subtracting (3.1.28) from it gives

$$Z = \frac{16Z_{e2}(h/2) - Z_{e2}(h)}{15} + O(h^6), \tag{3. 1.30}$$

define

$$Z_{e3}(h) = \frac{4^2 Z_{e2}(h/2) - Z_{e2}(h)}{4^2 - 1}. \tag{3. 1.31}$$

This process can be repeated n times, provided that the error form for the approximation of $Z_e(h)$ to Z can be expressed as

$$Z = Z_e(h) + \sum_{i=1}^{n-1} k_i h^{2i} + O\left(h^{2n}\right), \tag{3. 1.32}$$

where the constants k_i are independent of h. The $O(h^{2i})$ approximations are generated recursively by the formula

$$Z_{ei}(h) = \frac{4^{i-1} Z_{e(i-1)}(h/2) - Z_{e(i-1)}(h)}{4^{i-1} - 1} \tag{3. 1.33}$$

for $i = 2, 3, \cdots, n$.

The same outline will be used to obtain more accurate solutions for the extended systems; that is, the discretized form of the system (3.1.21) and the equation or equations which characterizes the singularity in the three previously mentioned cases, the simple turning points, bifurcation points and the cubic turning points. It was proven in chapter two, that the singular points are isolated solutions to the extended systems. This will allow us to generalize the above results to be used in such situations as was demonstrated in Attili and Shehadeh [5], Keller[10] and Stetter[18].

3.2 Block Elimination

Looking at the systems (3.1.6) to (3.1.8) in simple turning point case, (3.1.10) to (3.1.12) bifurcation point case and (3.1.15) to (3.1.19) in cubic turning point case, one realizes that F_x is the main block of the matrices involved. This gives rise to the idea of deflated block-elimination which requires one LU - factorization of F_x and then back-substitutions will only be needed to carry out the solutions of the previous systems. This is important since the LU - factorization requires $O(N^3)$ operations while the back substitution only requires $O(N^2)$ operations and to apply the extrapolation will mean 3-4 different values of step size h are needed. In addition to that, one can use the special structure of F_x; that is, tridiagonal matrix, which means the LU-factorization practicaly requires $O(N)$ operations. This will give rise to solving more and more systems with the minimum number of LU - factorizations. Not only this but we will show that even the calculation of $g_x, g_\lambda, g_\mu, \bar{g}_x, \bar{g}_\lambda$ and $\bar{g}_\mu$ will require the solution of systems with F_x being the main block.

Let us start first with the description of the block-elimination Chan[6] and Keller[10]. To solve

$$\begin{bmatrix} A & b \\ c^T & d \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ h \end{bmatrix}, \qquad (3.\ 2.1)$$

where A is $n \times n$ matrix, b a column vector, c^T a row vector and d a scalar. The following algorithm has the desirable property described earlier:

3.2.1 Algorithm

(1) Solve

(a) $A \cdot v = b$,

(b) $A \cdot w = f$.

(2) Compute $y = (h - c^T \cdot w)/(d - c^T \cdot v)$.

(3) Compute $x = w - yv$.

This can be seen by substituting y and x which are in step (2) and (3) respectively in left hand side of (3.2.1) to obtain

$$\begin{bmatrix} A & b \\ c^T & d \end{bmatrix} \cdot \begin{bmatrix} w - yv \\ (h - c^T w)/(d - c^T v) \end{bmatrix},$$

which after expanding equals to

$$\begin{bmatrix} A \cdot w - yA \cdot v + yb \\ c^T \cdot w - yc^T \cdot v + dy \end{bmatrix}$$

$$= \begin{bmatrix} A \cdot w - yA \cdot v + yA \cdot v \\ c^T \cdot w - c^T \cdot v \left(h - c^T \cdot w\right) / \left(d - c^T \cdot v\right) + d \left(h - c^T \cdot w\right) / \left(d - c^T \cdot v\right) \end{bmatrix}$$

$$= \begin{bmatrix} f \\ c^T \cdot w + \left(d - c^T \cdot v\right)\left(h - c^T \cdot w\right) / \left(d - c^T \cdot v\right) \end{bmatrix}$$

$$= \begin{bmatrix} f \\ c^T \cdot w + h - c^T \cdot w \end{bmatrix} = \begin{bmatrix} f \\ h \end{bmatrix},$$

which is the right hand side of (3.2.1).

A slight modification of the previous algorithm is needed to solve (3.1.7) which is of the form

$$(x, y) \begin{bmatrix} A & b \\ c^T & d \end{bmatrix} = (f, h), \tag{3. 2.2}$$

to solve the system we will have the following algorithm.

3.2.2 Algorithm

(1) Solve

(a) $v \cdot A = c^T$,

(b) $w \cdot A = f$.

(2) Compute $y = (h - b \cdot w)/(d - b \cdot v)$.

(3) Compute $x = w - yv$.

In step (1) of the above algorithm one practically solve $A^T \cdot v^T = c$ and $A^T \cdot w^T = f^T$ where the same LU-factorization of A is used.

Using Algorithm (3.2.1) and Algorithm (3.2.2), we will solve the systems (3.1.6), (3.1.15) and (3.1.7), (3.1.17) respectively. For example, to solve the system (3.1.7), by using Algorithm (3.2.2), we have

(1) Solve

(a) $F_z^T \cdot X_1^T = T$,

(b) $F_z^T \cdot X_2^T = 0$.

(2) Compute $g = (1 - R \cdot X_2)/(0 - R \cdot X_1)$.

(3) Compute $v = X_2 - g \cdot X_1$.

To solve the system (3.1.6) by using Algorithm (3.2.1), we have

(1) Solve

(a) $F_z \cdot X_1 = R$

(b) $F_z \cdot X_2 = 0$.

(2) Compute $g = (1 - T^T \cdot X_2)/(0 - T^T \cdot X_1)$.

(3) Compute $u = X_2 - gX_1$.

To solve (3.1.10) to (3.1.12), (3.1.16), (3.1.18) and (3.1.19), we have to expand step (1) in Algorithm (3.2.1) and Algorithm (3.2.2), since b and c^T consist of two column and two row vectors respectively. For example to solve the sytem (3.1.16) and with $\tilde{R} = (\tilde{R}_1, \tilde{R}_2)^T$, where $\tilde{R}_1$ is an $n \times 1$ vector and $\tilde{R}_2$ is a scalar, we will have the following Algorithm:

3.2.3 Algorithm

(1) Solve

(a) $F_z \cdot X_1 = F_\lambda$,

(b) $F_z \cdot X_2 = \tilde{R}_1$.

Compute $Y_{12} = (\tilde{R}_2 - g_z \cdot X_2)/(g_\lambda - g_z \cdot X_1)$,

and $Y_{11} = X_2 - Y_{12} X_1$.

Then $Y_1 = (Y_{11}, Y_{12})^T$.

(2) Solve

(a) $F_z \cdot \tilde{X}_1 = F_\lambda$.

(b) $F_z \cdot \tilde{X}_2 = 0$.

Compute $Y_{22} = \left(0 - g_z \cdot \tilde{X}_2\right) / (g_\lambda - g_z \cdot X_1)$,

and $Y_{21} = \tilde{X}_2 - Y_{22} \tilde{X}_1$.

Then $Y_2 = (Y_{21}, Y_{22})^T$.

(3) Compute

(a) $\tilde{g} = \left(1 - \tilde{T}^T \cdot Y_2\right) / \left(0 - \tilde{T}^T \cdot Y_1\right)$,

(b) $\tilde{v} = Y_2 - \tilde{g} Y_1$.

Theorem 3.2.1 *The system (3.1.16) is uniquely solvable and the solution is given by*

$$\tilde{g} = \left(1 - \tilde{T}^T \cdot Y_2\right) / \left(-\tilde{T}^T \cdot Y_1\right) \quad \text{and} \quad \tilde{v} = Y_2 - \tilde{g} Y_1$$

where Y_1 and Y_2 are as given in Algorithm (3.2.3).

Proof:

The unique solvability of the system was proven in theorem (2.3.1). To prove the second part consider

$$\begin{bmatrix} f_x & F_\lambda & \tilde{R}_1 \\ g_x & g_\lambda & \tilde{R}_2 \\ \tilde{T}^T & & 0 \end{bmatrix} \cdot \begin{bmatrix} Y_2 - \tilde{g} Y_1 \\ \tilde{g} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{pmatrix} F_x & F_\lambda \\ g_x & g_\lambda \end{pmatrix} (Y_2 - \tilde{g} Y_1) + \left(\tilde{R}_1, \tilde{R}_2\right)^T \tilde{g} \\ \tilde{T}^T \cdot (Y_2 - \tilde{g} Y_1) \end{bmatrix}$$

$$
= \begin{bmatrix} \begin{pmatrix} F_z & F_\lambda \\ g_z & g_\lambda \end{pmatrix} \begin{pmatrix} Y_{21} - \tilde{g}Y_{11} \\ Y_{22} - \tilde{g}Y_{12} \end{pmatrix} + \begin{pmatrix} \tilde{g}\cdot\tilde{R}_1 \\ \tilde{g}\cdot\tilde{R}_2 \end{pmatrix} \\ \tilde{T}^T\cdot Y_2 - \tilde{g}\tilde{T}^T\cdot Y_1 \end{bmatrix}
$$

$$
= \begin{bmatrix} F_z\cdot Y_{21} - \tilde{g}F_z\cdot Y_{11} + F_\lambda\cdot Y_{22} - \tilde{g}F_\lambda\cdot Y_{12} + \tilde{g}\tilde{R}_1 \\ g_z\cdot Y_{21} - \tilde{g}g_z\cdot Y_{11} + g_\lambda Y_{22} - \tilde{g}g_\lambda\cdot Y_{12} + \tilde{g}\tilde{R}_2 \\ \tilde{T}^T\cdot Y_2 - \tilde{T}^T\cdot Y_1\left(1-\tilde{T}^T\cdot Y_2\right)/\left(-\tilde{T}^T\cdot Y_1\right) \end{bmatrix}
$$

$$
= \begin{bmatrix} F_z\cdot\tilde{X}_2 - Y_{22}F_z\cdot\tilde{X}_1 - \tilde{g}F_z\cdot X_2 + \tilde{g}Y_{12}F_z\cdot X_1 + Y_{22}F_z\cdot\tilde{X}_1 - \tilde{g}Y_{12}F_z\cdot\tilde{X}_1 \\ +\tilde{g}F_z\cdot X_2 \\ g_z\cdot\tilde{X}_2 - Y_{22}g_z\cdot\tilde{X}_1 - \tilde{g}g_z\cdot X_2 + \tilde{g}Y_{12}g_z\cdot X_1 + g_\lambda Y_{22} - \tilde{g}g_\lambda Y_{12} + \tilde{g}\tilde{R}_2 \\ \tilde{T}^T\cdot Y_2 + 1 - \tilde{T}^T\cdot Y_2 \end{bmatrix}
$$

$$
= \begin{bmatrix} 0 \\ g_z\cdot\tilde{X}_2 + Y_{22}\left(g_\lambda - g_z\cdot\tilde{X}_1\right) - \tilde{g}g_z\cdot X_2 + \tilde{g}Y_{12}(g_z\cdot X_1 - g_\lambda) + \tilde{g}\tilde{R}_2 \\ 1 \end{bmatrix}
$$

$$
= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},
$$

which is the right hand side of the system (3.1.16).

Note that X_1 and $\tilde{X}_1$ in part (a) of steps 1 and 2 of the algorithm are the same since the right hand side is the same in both systems. Thus there is no need to repeat the back-substitution to find $\tilde{X}_1$ once X_1 is found.

To solve the system (3.1.17), that is

$$\left(\tilde{u}^T, \tilde{g}\right) \begin{bmatrix} F_x & F_\lambda & \tilde{R} \\ g_x & g_\lambda & \\ \tilde{T}^T & & 0 \end{bmatrix} = [0, 1],$$

with $\tilde{T}^T = \left(\tilde{T}_1^T, \tilde{T}_2\right)$ where $\tilde{T}_1^T$ is an $n \times 1$ vector and $\tilde{T}_2$ is a scalar, we will have the following algorithm.

3.2.4 Algorithm

(1) Solve

(a) $F_x^T \cdot v_1^T = g_x^T$,

(b) $F_x^T \cdot v_2^T = \tilde{T}_1$.

Compute $w_{12} = \left(\tilde{T}_2 - F_\lambda^T \cdot v_2^T\right) / \left(g_\lambda - F_\lambda^T \cdot v_1^T\right)$,

and $w_{11} = v_2 - w_{12} v_1$.

Then $w_1 = (w_{11}, w_{12})^T$.

(2) Solve

(a) $F_x^T \cdot \tilde{v}_1^T = g_x^T$,

(b) $F_x^T \cdot \tilde{v}_2^T = 0$.

Compute $w_{22} = \left(0 - F_\lambda^T \cdot \tilde{v}_2^T\right)\left(g_\lambda - F_\lambda^T \cdot v_1^T\right)$,

and $w_{21} = \tilde{v}_2 - w_{22}\tilde{v}_1$

Then $w_2 = (w_{21}, w_{22})^T$.

(3) Compute $\tilde{u} = w_2 - \tilde{g}w_1$.

Note that v_1 and $\tilde{v}_1$ are identical since the system to be solved in both cases is the same.

It should be noted here that block-elimination may be unstable since F_x will be nearly singular. When numerically tested, however, it was fairly reliable and only fails when F_x is very singular, at which point the accuracy is usually high enough to stop the iterations. Still to avoid the diffculties which might arise when F_x is singular or nearly singular, consider the expanded form of (3.2.1); that is,

$$Ax + by = f$$

$$c^T x + dy = h. \tag{3. 2.3}$$

Since ψ_0^T is a left null vector of A, we will have

$$y = \psi_0^T f / \psi_0^T b. \tag{3. 2.4}$$

where $u^T = \psi_0^T$ at the singular point. Here (3.2.4) is well defined since $\psi_0^T b \neq 0 (u^T R = 1)$. If (3.2.4) is substituted in (3.2.3), we obtain

$$Ax = f - b(\psi_0^T f / \psi_0^T b). \tag{3. 2.5}$$

The right hand side of (3.2.5) is in the range of A which means (3.2.5) is solvable and the general solution is

$$x = x_p + \gamma_0 \phi_0, \tag{3. 2.6}$$

where x_p is any particular solution and $\gamma_0 \in R$. With this we will have the following theorem.

Theorem 3.2.2 *The system (3.2.3) has a unique solution given by*

$$y = \psi_0^T f / \psi_0^T b, \qquad x = x_p + \gamma_0 \phi_0$$

where γ_0 is given by

$$\gamma_0 = (h - dy - c^T x_p)/c^T \phi_0$$

and ψ_0 and ϕ_0 are the left and right null vectors of $A = F_x$ at the singular point if and only if $\psi_0^T b \neq 0$ and $c^T \phi_0 \neq 0$.

Proof:

The proof is straight-forward and can be seen from equations (3.2.4) and (3.2.6).

Note that the two conditions mentioned in the theorem, namely, $\psi_0^T b \neq 0$ and $c^T \phi_0 \neq 0$ are guaranteed for our situation since $u^T R = 1$ and $T^T v = 1$.

3.3 The Partials of g and $\tilde{g}$

To carry out the Newton's iterations, one also needs the partials of g with respect to z and λ in the simple turning point and simple bifurcation point cases and the partials of g and $\tilde{g}$ with respect to z, λ and μ in the cubic turning point case. The main difficulty in doing so is that the gradient of g and $\tilde{g}$, however they are defined, depends on second derivatives of F which may be costly to evaluate specially if finite-differences are to be used. To calculate the gradient of g, with respect to $z_i; i = 1, 2, \cdots N$, one can use the formulas

$$\begin{aligned} g_{zi} &= \frac{g(z_1, z_2, \cdots, z_i + \epsilon, \cdots, z_N, \lambda) - g(z_1, z_2, \cdots, z_N, \lambda)}{\epsilon} \\ &= -u^T \left(\frac{F_z(z_1, z_2, \cdots, z_i + \epsilon, \cdots, z_N, \lambda) - F_z(z_1, z_2, \cdots, z_N, \lambda)}{\epsilon} \right) V, \end{aligned}$$

also

$$g_\lambda = -u^T \left(F_z(z_1, z_2, \cdots, z_N, \lambda + \epsilon) V - F_z(z_1, z_2, \cdots, z_N, \lambda) V \right) / \epsilon.$$

The above two formulas have the disadvantage of recalculating the Jacobian $(N+1)$ times at each iteration which will be very costly. To reduce the amount of work used one can use the approximation

$$g_z \approx \frac{-u^T (F_z(z + \epsilon V, \lambda) - F_z(z, \lambda))}{\epsilon}$$

in the turning point cases and

$$(g_z)_i \approx \frac{-u^T(F_z(z+\epsilon V_i,\lambda)-F_z(z,\lambda))}{\epsilon}; i=1,2$$

in the bifurcation point case. It has the advantage that it only requires one extra evaluation of the Jacobian at each iteration.

Another method to compute the partials of g and $\tilde{g}$ is as follows.

Consider

$$g = -u^T F_z v, \tag{3. 3.1}$$

as defined in Lemma (2.1.1). Considering the special case $R = F_\lambda$ (F_λ is not in the range of F_z) and differentiating g with respect to z, we obtain

$$g_z = -u_z^T F_z v - u^T F_{zz} v - u^T F_z v_z \tag{3. 3.2}$$

From the determining systems of g, u and v; that is, (3.1.15) and (3.1.17), we will have

$$F_z v = -gF_\lambda,\ \ u^T F_z = -gT^T,\ \ T^T v_z = 0 \quad u^T F_{z\lambda} = -u_z^T F_\lambda$$

which when substituted in (3.3.2) lead to

$$g_z = -u^T F_{z\lambda} g - u^T F_{zz} v. \tag{3. 3.3}$$

Using the same approach one obtains,

$$g_\lambda = -u^T F_{\lambda\lambda} g - u^T F_{z\lambda} v. \tag{3. 3.4}$$

Combining (3.3.3) and (3.3.4) we obtain

$$(g_z, g_\lambda) = \left(-u^T F_{zz} - u^T F_{z\lambda}, -u^T F_{z\lambda} - u^T F_{\lambda\lambda}\right) \begin{pmatrix} v \\ g \end{pmatrix} \tag{3. 3.5}$$

or equivalently

$$\nabla_{z\lambda}\, g = -u^T(\nabla^2_{z\lambda} F) \begin{pmatrix} v \\ g \end{pmatrix} \tag{3. 3.6}$$

.

Equation (3.3.6) gives away of evluating g_z and g_λ in the simple and cubic turning point cases. In the simple bifurcation point case, v and g are composed of two components of each and so (3.3.6) becomes

$$\nabla_{z\lambda} g_i = -u^T \left(\nabla^2_{z\lambda F}\right) \begin{pmatrix} v_i \\ g_i \end{pmatrix}, \quad i = 1,2$$

To evaluate the other partials of g and $\tilde{g}$ we consider the special case of having $R = F_\lambda$ in the determining system for g and $\tilde{R} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tilde{v} = \begin{pmatrix} w \\ \delta \end{pmatrix}$ and $\tilde{T}^T = (T^T, 0)$ in the determining system for $\tilde{g}$; that is, respectively (3.1.15) and (3.1.16) becomes

i)

$$F_z v + F_\lambda g = 0$$

ii)

$$T^T v = 0 \tag{3. 3.7}$$

and

i)

$$F_x w + \delta F_\lambda = 0$$

ii)

$$g_x w + \delta g_\lambda + \tilde{g} = 0$$

iii)

$$T^T w = 1 \qquad (3.\ 3.8)$$

Comparing (3.3.7) and (3.3.8), we conclude that $v = w, g = \delta$ and $\tilde{g} = -g_x v - g g_\lambda$. This will imply that

i)

$$\tilde{g}_x = -g_x v_x - g_{xx} v - g_x g_\lambda - g g_{x\lambda}$$

ii)

$$\tilde{g}_\lambda = -g_x v_\lambda - g_{x\lambda} v - g_\lambda g_\lambda - g g_{\lambda\lambda}$$

iii)

$$\tilde{g}_\mu = -g_x v_\mu - g_{x\mu} v - g_\mu g_\lambda - g g_{\lambda\mu} \qquad (3.\ 3.9)$$

Let us consider the computation of (3.3.9i) only since the other two formulas will be done in a similar way. To do so one needs v_x, g_{xx} and $g_{x\lambda}$ or an approximation to them since we already computed g_x, v, g_λ and g.

To compute v_z, differentiate (3.3.7) with respect to z, to obtain

$$F_{zz}v + F_z v_z + F_{\lambda z}g + F_\lambda g_z = 0$$

and

$$T^T v_z = 0$$

or equivalently

$$\begin{pmatrix} F_z & F_\lambda \\ T^T & 0 \end{pmatrix} \begin{pmatrix} v_z \\ g_z \end{pmatrix} = \begin{pmatrix} -F_{zz}v - gF_{z\lambda} \\ 0 \end{pmatrix}. \qquad (3.\ 3.10)$$

The matrix on the left hand side of (3.3.10) is the same one used in (3.3.7), which means we can compute the value of g_z while solving for v_z. An advantage of using the above system in that block-elimination can be used to obtain g_z and v_z. Similar arguments will be needed and similar systems are to be solved to calculate all other partials of g and $\tilde{g}$ in which we make use of the LU-factorization of F_z. This means that 6-systems which involve the same matrix in (3.3.10) are needed to calculate the partials of g and $\tilde{g}$. As a result, one can solve 10 systems each iteration using one LU-factorization of the main block F_z. Which means the amount of work required will be reduced due to the difference in the order of operations between the LU-factorization and the back substituation.

3.4 Choice of R and T

The simplist possible choice of R, which is a column vector, is to choose it from the bases vector e_i, where e_i is the zero vector of size N with 1 in the i^{th} component. The i^{th} component is chosen in a way that makes (F_z, R) of full rank as was pointed earlier. To choose T which is a normalization on the null vector of F_z, one can choose it as $(I, 0)$ or $(0, I)$ where I is an identity matrix of proper size. Such choice is one of the simplist.

It we choose R to be $R = (r_1, r_2, \cdots, r_N)$ and since R is not in the range of F_z we will have $u^T R \neq 0$. But $u^T R = \sum_{i=1}^{N} u_i^T r_i$ which may be interpreted as a quadrature approximation to $\int_a^b \psi_0^T R dx$, which is assumed not zero.

Similarly for T; that is, if we choose $T^T = (t^T, t_2^T, \cdots, t_N^T)$ and $v = (v_1, v_2, \cdots, v_N)$. Now since $T^T v \neq 0 = \sum_{i=1}^{N} t_i^T v_i$ which also be interpreted as a quadrature approximation of $\int_a^b T^T \cdot \phi_0 dx$, which is assumed not zero.

One further way to choose R is to notice that (F_z, R) should have full rank. Which implies that R is not in the range of F_z. But the range of F_z is equal the null space of $(F_z^T)^\perp$; that is, $R(F_z) = N(F_z^T)^\perp$, this implies that if we choose $R \in N(F_z^T)$, then R is out of $R(F_z)$. Thus we could choose R such that $F_z^T R = \epsilon$, where ϵ is a very small number. One might make use of the LU factorization of F_z in finding R and use block elimination techniques to solve the systems involved.

Now writing F_z in factorization form; that is,

$$F_z = LU, \tag{3. 4.1}$$

since F_z is nearly singular, then the upper triangular matrix U in (3.4.1) will have the form

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & \cdots & & u_{1N} \\ 0 & u_{22} & \cdots & \cdots & & u_{2N} \\ \vdots & & & & & \vdots \\ \vdots & & & & u_{N-1,N-1} & u_{N-1,N} \\ 0 & \cdots & & & 0 & \epsilon \end{bmatrix}.$$

Let e_n be the nth coordinate vector. Now by defining

$$R = (L^{-1})^T e_n,$$

we will have

$$R^T F_z = e_n^T L^{-1} LU = e_n^T U = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \epsilon \end{pmatrix}.$$

Thus to find R, we will solve the linear system $L^T R = e_n$. This can be computed by back substitution.

Chapter 4

NUMERICAL EXAMPLES

In this chapter, we present three examples of the three different types of singularities which were discussed before, namely, the simple turning points, cubic turning points and bifurcation points. Two different approaches will be used to solve the linear systems involved. The first will be using LU-factorization with block elimination and the second will be LU-factorization without block elimination. To get more accurate approximations we will apply Richardson extrapolation. The CPU time for each method will be calculated to show the advantages of using block elimination.

The following examples were used for numerical experimentation

(1) We solved the one-dimensional nonlinear problem

$$y'' + \lambda e^{y} = 0 \text{ on the interval } [0, 1]$$

With the boundary conditions $y(1) = y(0) = 0$ which has a simple turning point at the critical parameter, $\lambda = 3.513807$. With the initial guess $\lambda = 3.4$ and using the extended system (3.1.5) The results for the fifth iteration with $h = \frac{1}{3}, \frac{1}{6}, \frac{1}{12}$ and $\frac{1}{24}$ using block elimination are given in Table 1.

Table 1

h	g	λ
$\frac{1}{3}$	0.143051E-05	3.31092
$\frac{1}{6}$	-0.722452E-05	3.46261
$\frac{1}{12}$	0.560958E-05	3.50110
$\frac{1}{24}$	0.471423E-05	3.51062

Using the results in Table 1, we applied the Richardson extrapolation and obtained $\lambda = 3.51378$ where the solution we are after is $\lambda = 3.5138307$. The results of the application of the extrapolation are given in Table 2.

Table 2.

h	Number of Extrapolation			
	0	1	2	3
$\frac{1}{3}$	3.31092			
		3.51318		
$\frac{1}{6}$	3.46261		3.51398	
		3.51393		3.51378
$\frac{1}{12}$	3.50110		3.51378	
		3.51379		
$\frac{1}{24}$	3.51062			

It is clear from Table 1 that $g \to 0$ for the various values of h as expected. The CPU time is equal to 0.040 sec.

Repeating the same calculations for $h = \frac{1}{3}, \frac{1}{6}, \frac{1}{12}$ and $\frac{1}{4}$ but without using the block-elimination this time. The results are given in Table 3.

Table 3

h	g	λ
$\frac{1}{3}$	0.1172142E-03	3.31091
$\frac{1}{6}$	.395983E-05	3.46261
$\frac{1}{12}$	.171911E-05	3.50110
$\frac{1}{24}$	-.404437E-05	3.51062

Again applying Richardson extrapolation on the approximate values of λ_h the results are given in Table 4.

Table 4.

h	Number of Extrapolation			
	0	1	2	3
$\frac{1}{3}$	3.31092			
		3.51318		
$\frac{1}{6}$	3.46261		3.51398	
		3.51393		3.51378
$\frac{1}{12}$	3.50110		3.51378	
		3.51379		
$\frac{1}{24}$	3.51062			

The CPU time is equal to 0.240 sec. It is clear that the CPU time used with block elimination is $\frac{1}{6}$ of the time used without block elimination.

We also solved the system with $h = \frac{1}{50}$ using block elimination. With the same initial guess $\lambda = 3.4$, we obtain the solution $\lambda = 3.51295$ and $g = .98533E-04$. The CPU time was 0.190 sec. This shows that with such a large system, we were not able to obtain the same accuracy as was obtained in Table 3, also the time required is 4 times that needed with Richardson extrapolation.

(2) We consider the two parameter problem

$$y'' + \lambda \exp\left(\frac{y}{1+\mu y}\right) = 0, \qquad\qquad y(0) = y(1) = 0,$$

which has a cubic turning point at $\lambda = 5.22949$ and $\mu = 0.2457804$, See Attili [1]. With the initial guesses $\lambda = 5.0$ and $\mu = 0.2$, and using the extended system (3.1.14), the results after four iterations with $h = \frac{1}{3}, \frac{1}{6}, \frac{1}{12}$ and $\frac{1}{24}$ using block elimination are given in Table 5.

Table 5

h	g	$\bar{g}$	λ	μ
$\frac{1}{3}$	.107766E-03	.419806E-02	4.87464	.250123
$\frac{1}{6}$	.201616E-05	0.133458E-04	5.14327	.246672
$\frac{1}{12}$	.230550E-04	.591536E-05	5.20793	.245981
$\frac{1}{24}$	-.454786E-04	.603387E-04	5.22398	.245823

Applying Richardson extrapolation for the results in Table 5 we obtain $\lambda = 5.22931$ and $\mu = .245772$, where the solution we are after is $\lambda = 5.229494$ and $\mu = .2457804$ at the cubic turning point. The results of the application of the extrapolation are given in Table 6 & 7.

Table 6.

h	Number of Extrapolation			
	0	1	2	3
$\frac{1}{3}$	4.87464			
		3.23282		
$\frac{1}{6}$	5.14327		5.22926	
		5.22948		5.22931
$\frac{1}{12}$	5.20793		5.22931	
		5.22932		
$\frac{1}{24}$	5.22397			

Table 7.

h	Number of Extrapolation			
	0	1	2	3
$\frac{1}{3}$	.250123			
		.245522		
$\frac{1}{6}$	.246672		.245765	
		.245750		.245772
$\frac{1}{12}$	.245981		.245772	
		.245771		
$\frac{1}{24}$	.245823			

It is clear from Table 5 that $g \to 0$ and $\tilde{g} \to 0$ for various values of h as expected. The CPU time required is 0.160 Sec.

Repeating the same calculations for $h = \frac{1}{3}, \frac{1}{6}, \frac{1}{12}$ and $\frac{1}{24}$ but without using the block elimination. The results are given in Table 8.

Table 8

h	g	$\bar{g}$	λ	μ
$\frac{1}{3}$	.939375E-04	.230773E-02	4.87272	.250034
$\frac{1}{6}$	.251079E-04	.365602E-04	5.14327	.246672
$\frac{1}{12}$	.7947520E-04	.305270E-04	5.20797	.245983
$\frac{1}{24}$	-.213198E-05	.772805E-06	5.22404	.245826

Applying Richardson extrapolation on the approximate values of λ_h we obtained $\lambda = 5.22938$ and $\mu = .245776$ The CPU time is 1.35 sec. Then the CPU time used with block elimination is less than $\frac{1}{8}$ the time used without block elimination.

(3) We consider the two point boundary value problem

$$x''(s) - p(\lambda)X''(s) + \pi^2\lambda f(x(s) - p(\lambda)X(s)) = 0$$

$x(0) = x(1) = 0$,
where $p(\lambda) = \lambda^4 \exp(-\lambda/2), f(z) = z^2 + z$ and $X(s) = s(1-s)\exp(s)$, which has a simple bifurcation point at $\lambda = 1$. With the initial guess $\lambda = .9$ and $\gamma = .01$. With $h = \frac{1}{36}$ the result for the fifth iteration using block elimination is $\lambda = 1.00025$. As expected $\gamma \to 0$ and both components of g did the same. The CPU time required is .08 sec.

Repeating the same calculations for the same h but without using block elimination, the results we obtained were very similar to the ones with block elimination. The CPU time required is .83 sec. This means that the CPU time used with block elimination is less than $\frac{1}{10}$ the time used without block elimination.

Bibliography

[1] Attili, B. S., *Multiple Shooting and Calculations of Some Types of Singularities in B.V.P.'s*, Inter. J. Computer Math. 32(1990), 97-111.

[2] Attili, B. S., *Characterization and Computation of Bifurcation Points with Corank 2*, Proc. 13th IMACS World Congress on Computation and Applied Math., Dublin, 1991.

[3] Attili, B. S., *Numerical Computation of Symmetry Breaking Bifurcation Point*, J. of the Aust. Math. Soci Series B; Applied Math. (to appear)

[4] Attili, B. S., *A Direct Method for the Characterization and Computation of Bifurcation Points with Corank 2*, Computing 47(1992), (to appear).

[5] Attili, B. S. and Shehadeh Y., *Existence of asymptotic expansions of the extended systems used in the computation of singular points*, (under preparation).

[6] Chan, T. F. , *Deflation Techniques and Block-Elimination Algorithms for Solving Bordered Singular Systems*, Technical Report # 226 March 10, 1982 Computer Science Department, Yale Univ., Yale Station, New Haven, CT 06520.

[7] Greiwank, A and Reddien, G. W., *Characterization and Computation of Generalized Turning Points*, SIAM J. Numer. Anal. 21(1984) 186-196.

[8] Ioos, Joseph, *Elementary Stability and Bifurcation Theory*, Second Edition, Springer-Verlag, (1990).

[9] Jepson, A. and Spence, A., *Folds in Solutions of Two Parameter Systems and their Calculation*, Part I, SIAM J. Numer. Anal. 22 (1985), 347-369.

[10] Keller, H. B. , *Accurate Difference Methods for Nonlinear Two-Point Boundary Value Problems*, SIAM J. Numer. Anal 2(1974) 305-320.

[11] Keller, H. B., *Numerical Solutions of Bifurcation and Nonlinear Eigenvalue Problems in Application of Bifurcation Theory*, P.H. Rabinowitz, Ed. Academic Press, New York (1977) 359-389.

[12] Kikuchi, F., *An Iterative Finite Element Scheme for Bifurcation Analysis of Semi- Linear Elliptic Equations*, Institute of Space and

Aeronautical Science, University of Tokyo, Report No.542 (Vol.41, No.6),(1976).

[13] Langford, W. F. , *A Shooting Algorithm for the Best Least Squares Solutions of Two Point Boundary Value Problems*, SIAM J. Numer. Anal. 14(1977), 527-542.

[14] Moore and Spence, *The Calculation of Turning Points of Nonlinear Equations*, SIAM J. Numer. Anal. 17(1980), 567-576.

[15] Rheinboldt, W.C., *Numerical Methods for a Class of Finite Dimensional Bifurcation Problems*, SIAM J. Numer. Anal, 15(1978), 1-11.

[16] Stakgold, I. ,*Branching Solutions of Nonlinear Equations*, SIAM Review, 13(1971), 283-332.

[17] Spence, A. and Werner, B. , *Nonsimple Turning Points and Cusps*, SIAM J. Numer. Anal. 2(1982), 413-427.

[18] Stetter, H., *Asymptotic Expansions for the Error Discretization Algorithms for Nonlinear Functional Equations*, Numer. Amath., 7 (1965), 18-31.

[19] Weiss, R. , *The Convergence of Shooting Methods*, BIT 13(1973), 270-475.